REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	New Reprint	-

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Self-Assembled Multi-Component Catenanes: The Effect of Multivalency and Cooperativity on Structure and Stability	
	5b. GRANT NUMBER
	W911NF-04-D-0004
	5c. PROGRAM ELEMENT NUMBER
	611102

6. AUTHORS	5d. PROJECT NUMBER
Stephen J. Lee, Marcey L. Waters, Michel R. Gagné, Mee-Kyung Chung	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
University of North Carolina - Chapel Hill	
Office of Sponsored Research	
104 Airport Drive, Suite 2200, CB #1350	
Chapel Hill, NC 27599 -1350	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	ARO
U.S. Army Research Office	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
P.O. Box 12211	
Research Triangle Park, NC 27709-2211	58118-CH-SR.3

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Using dynamic combinatorial chemistry (DCC), mixtures of dipeptide monomers were combined to probe how the structural elements of a family of self-assembled [2]-catenanes affect their equilibrium stability versus competing non-catenated structures. Of particular interest were experiments to target the role of CH-? interactions, inter-ring hydrogen bonds and ?-turn types on [2]-catenane energetics. The non-variant core of the [2]-catenane was shown to only adopt type II? and type VIII turns at the ?-2 and ?-4 positions, respectively. Monomers were designed to

15. SUBJECT TERMS

molecular recognition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Stephen Lee
UU	UU	UU	UU		19b. TELEPHONE NUMBER
					919-549-4365

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Self-Assembled Multi-Component Catenanes: The Effect of Multivalency and Cooperativity on Structure and Stability

ABSTRACT

Using dynamic combinatorial chemistry (DCC), mixtures of dipeptide monomers were combined to probe how the structural elements of a family of self-assembled [2]-catenanes affect their equilibrium stability versus competing non-catenated structures. Of particular interest were experiments to target the role of CH-? interactions, inter-ring hydrogen bonds and ?-turn types on [2]-catenane energetics. The non-variant core of the [2]-catenane was shown to only adopt type II? and type VIII turns at the ?-2 and ?-4 positions, respectively. Monomers were designed to delineate how these factors contribute to [2]-catenane equilibrium speciation/stability. Dipeptide turn adaptation studies, including three-component dynamic self-assembly experiments, suggested that stability losses are localized to the mutated sites, and that the turn types for the core ?-2 and ?-4 positions, type II? and type VIII, respectively, cannot be modified. Mutagenesis studies on the core Aib residue involved in a seemingly key CH-?-CH sandwich reported on how CH-? interactions and inter-ring hydrogen bonds affect stability. The interacting methyl group of Aib could be replaced with a range of alkyl and aryl substituents with monotonic affects on stability, though polar heteroatoms were disproportionately destabilizing. The importance of a key cross-ring H-bond was also probed by examining an Aib for L-Pro variant. Inductive affects and the effect of CH donor multiplicity on the core proline-? interaction also demonstrated that electronegative substituents and the number of CH donors can enhance the effectiveness of a CH-? interaction. These data were interpreted using a cooperative binding model wherein multiple non-covalent interactions create a web of interdependent interactions. In some cases, changes to a component of the web leads to compensating effects in the linked interactions, while in others, the perturbations create a cascade of destabilizing interactions that lead to disproportionate losses in stability.

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Continuation for Block 13

ARO Report Number   58118.3-CH-SR
Self-Assembled Multi-Component Catenanes: T   ...

Block 13: Supplementary Note

Self-Assembled Multi-Component Catenanes: The Effect of Multivalency and Cooperativity on Structure and Stability
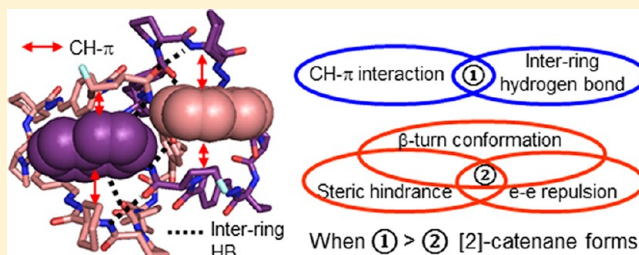
Mee-Kyung Chung,[†] Stephen J. Lee,[‡] Marcey L. Waters,*,[†] and Michel R. Gagné*,[†]

[†]Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3290, United States
[‡]U.S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709, United States

S *Supporting Information*

ABSTRACT: Using dynamic combinatorial chemistry, mixtures of dipeptide monomers were combined to probe how the structural elements of a family of self-assembled [2]-catenanes affect their equilibrium stability versus competing non-catenated structures. Of particular interest were experiments to target the effects of CH−π interactions, inter-ring hydrogen bonds, and β-turn types on [2]-catenane energetics. The non-variant core of the [2]-catenane was shown only to adopt type II′ and type VIII turns at the β-2 and β-4 positions, respectively. Monomers were designed to delineate how these factors contribute to [2]-catenane equilibrium speciation/stability. Dipeptide turn adaptation studies, including three-component dynamic self-assembly experiments, suggested that stability losses are localized to the mutated sites, and that the turn types for the core β-2 and β-4 positions, type II′ and type VIII, respectively, cannot be modified. Mutagenesis studies on the core Aib residue involved in a seemingly key CH−π−CH sandwich reported on how CH−π interactions and inter-ring hydrogen bonds affect stability. The interacting methyl group of Aib could be replaced with a range of alkyl and aryl substituents with monotonic affects on stability, though polar heteroatoms were disproportionately destabilizing. The importance of a key cross-ring H-bond was also probed by examining an Aib for L-Pro variant. Inductive affects and the effect of CH donor multiplicity on the core proline−π interaction also demonstrated that electronegative substituents and the number of CH donors can enhance the effectiveness of a CH−π interaction. These data were interpreted using a cooperative binding model wherein multiple noncovalent interactions create a web of interdependent interactions. In some cases, changes to a component of the web lead to compensating effects in the linked interactions, while in others, the perturbations create a cascade of destabilizing interactions that lead to disproportionate losses in stability.

■ INTRODUCTION

Multivalency and cooperative binding are phenomena that occur throughout the genre of molecular recognition, though the latter is principally observed in the biological context.[1] While many molecular systems participate in multivalency, true cooperativity, i.e., binding strengths that are greater than the sum of the parts, is rare for synthetic systems.[2] One contributor to the chasm between synthetic and biological systems is the inherently small size and the lack of complexity of synthetic systems that tend to be amenable to study by traditional spectroscopic and/or thermochemical means. The process of "designing" a molecular recognition system additionally creates a bias toward the molecular interactions that were used to create the structure's desired recognition properties. This heavy hand overwhelms the subtle factors that could conspire to create cooperativity and additionally distances the results from the biological context where they are most often observed. The study of multivalency and cooperative binding is therefore challenged by the constraints of being "designed" and the technical difficulties of studying/manipulating complex (large) systems known to display cooperativity. The generation of spontaneously assembled systems held together with multiple,

easily varied interactions creates a new tool for the study of multivalency and cooperative binding in complex systems.

In our previous paper,[3] we described a phenomenon wherein mixtures of dipeptides self-assemble under dynamic conditions into stereochemically and constitutionally precise [2]-catenanes, wherein one 56-membered macrocyclic ring interlocks with a second ring to give highly ordered structures that are intermediate in size and complexity between typical molecular recognition and biological systems. Despite their large size (>3 kDa), these compounds could be precisely characterized by high-resolution techniques that include X-ray crystallography, NMR, and mass spectrometry. These tools revealed that the [2]-catenanes are assembled from a complex web of molecular interactions that include both intra- and inter-macrocycle hydrogen bonds, a broad array of edge-to-face and slipped face-to-face π−π interactions, and seemingly critical CH−π interactions. Other defining structural features include the utilization of β-turns, which, at least in the core, are highly conserved and adopt biomimetic turn types (β-2 is type II′ and

β-4 is type VIII). Each of these non-covalent interactions and the β-turn secondary structures are common elements in protein structures.

In the [2]-catenanes characterized in our previous paper,[3] one invariant region of the structure is the "core", where the two conserved aryl glycine units seed a network of CH−π interactions and inter-macrocycle hydrogen bonds. The absolute requirement for an aryl glycine-containing monomer in the dynamic self-assembly reaction (DSA), and the observation that each arylGly unit engages in CH−π interactions on both π−faces, led to the hypothesis that these interactions, expressed through the CH−π−CH sandwich, were acting to template the synthesis and stabilize the catenane.

Relatively weak CH−π interactions are not well studied, despite their ubiquity and influence over important chemical phenomena, which include small-molecule conformation,[4] chiral discrimination,[5] assembly of host−guest complexes,[4a,6] molecular recognition,[7] selectivity of organic reactions,[6a,8] crystal packing,[6a,9] bioconjugates,[4] and protein−protein interactions involving polyproline helices.[10] The well-characterized nature of the catenane and the modular, thermodynamically controlled self-assembly/synthesis protocol suggested an experimental strategy for probing the importance and the properties of the CH−π interactions present in the self-assembled [2]-catenanes. The desire to examine the CH−π interaction in a complex setting was stimulated by the constraints of small-molecule studies, which can measure the strength of the CH−π interaction in isolation, and computational studies, which can probe the CH−π interaction in simple systems. Studies capable of manipulating this interaction in a complex, multivalent scheme that bridge the gap to the biological context are scant.

With the goal of delineating a structure/stability relationship, we report the outcome of our investigations to manipulate the [2]-catenane's defining CH−π interactions and subsequently its core β-turns. The previous paper[3] provided a structural blueprint for designing modified monomers to perturb the strength of the CH−π interaction, as determined from the relative stability of the folded [2]-catenane versus competing non-catenated species. The nature of these experiments required careful controls to support the notion that the examined subtle perturbations were principally affecting the highly structured and folded catenanes rather than the non-catenated cyclic oligomers.

These mutation studies indicate that a minimum number of CH−π interactions are necessary, *but not sufficient* for the self-assembly of these catenanes. Mutations in the CH−π participating core β-turns (β-2 and β-4) revealed that the conserved turn structures were critical to the stability of the [2]-catenanes. Structure−stability correlations showed the CH−π interactions to be part of an interconnected network of non-covalent interactions that encompass the entire structure. Semi-quantitative measurements of stability indicate that even small structural perturbations can become magnified or attenuated in this network. For example, the excision of two CH_3 groups involved in CH−π interactions causes a >1000-fold drop in catenane stability. As will be discussed, the source of this loss is a combination of CH−π loss, a change in β-turn conformational preference, and a concomitant loss of highly conserved inter-ring hydrogen bonds. In other mutations, the expected drop in stability was attenuated, and metrical evidence points to compensating structural reorganizations which restore some of the lost stability. These experiments thus enable one to examine multivalency and perhaps even cooperativity in a regime of significant structural complexity. As such, this system enables one to begin revealing how the nature of molecular interactions can be modulated by other structural elements.

■ EXPERIMENTAL SECTION

General Methods. Chemicals were purchased from Aldrich, Fisher Scientific, and Chem-Impex International, Inc. and used as received. ^{1}H and ^{13}C NMR spectra of monomers were recorded on a Bruker DRX 500 spectrometer and processed using TOPSPIN Bruker NMR software (V. 3.0). High-resolution mass spectra were obtained on an Agilent Accurate LC-TOF mass spectrometer (Agilent Series 6220) operating in positive-ion mode with an electrospray ionization source (fragmentor = 175 V). The data were analyzed using Agilent MassHunter Workstation Software, Qualitative Analysis (V. B.02.00). HPLC analysis was performed on a Hewlett-Packard Series 1100 instrument, using a Halo-C18 column (4.6 × 150 mm, 2.7 μm) with gradient elution (methanol/water) at a flow rate of 0.45 mL/min and at 55 °C. The injection volume for a 5 mM dynamic self-assembly (DSA) reaction was typically 3.5 μL. UV absorbance chromatograms were recorded at wavelengths of 220 and 289 nm. The data were analyzed using the Agilent Chemstation software. Accurate Mass LC-TOF analyses of DSA reactions were performed on an Agilent Series 1200 LC instrument, using a Halo-C18 column (2.1 × 50 mm, 2.7 μm) with gradient elution (methanol/water containing 0.1% formic acid) at a flow rate of 0.5 mL/min and at 50 °C. The eluent was analyzed by an Accurate Mass LC-TOF mass spectrometer (Agilent Series 6220) in positive-ion mode with an electrospray ionization source (fragmentor = 375 V). The data were analyzed using Agilent MassHunter Workstation Software, Qualitative Analysis (V. B.02.00).

Self-Assembly of [2]-Catenanes for HPLC and LC-TOF Analyses. Five millimolar DSA solutions were prepared on a 1 mL scale. D- or D,L-**1x** (5 μmol) and L,L-^{y}**2** or L,L-^{y}**2^{Ph}** (5 μmol) were separately dissolved in 1 mL of a mixture of MeCN and CHCl$_3$ (1:3), and 0.5 mL of the D- or D,L-**1x** solution was mixed with 0.5 mL of the L,L-^{y}**2** or L,L-^{y}**2^{Ph}** solution and trifluoroacetic acid (50 equiv). The solutions were allowed to sit for several days until a steady state was reached (typically 7−10 days) prior to LC analyses.

Speciation and Quantification of [2]-Catenanes by HPLC and LC-TOF. The speciation and characterization of [2]-catenanes in the DSA solution were performed by HPLC and LC-TOF analyses. For HPLC analyses, 3.5 μL of DSA solution was injected onto a Halo-C18 column (4.6 × 150 mm, 2.7 μm) at a flow rate of 0.45 mL/min at 55 °C. To remove solvent and TFA, a short linear gradient (35−46% MeOH/H$_2$O over 4 min) was used. A second gradient composed of two linear gradients (46−66% MeOH/H$_2$O over 26 min and consecutive 66% MeOH/H$_2$O to 100% MeOH over 12 min) was used to separate non-catenane species, and then 100% MeOH was subsequently used for an additional 12 min for [2]-catenane speciation. The [2]-catenanes in the DSA solution were quantified from their UV trace areas (289 nm) and compared to a calibration curve generated from known concentrations of several isolated catenanes and their corresponding UV trace areas (289 nm). The [2]-catenanes in the DSA solution were identified by LC-TOF analyses in positive-ion mode with an electrospray ionization source (fragmentor = 375 V, Vcap = 3500 V, mass range = 100−3200 m/z). A 1.0 μL DSA solution was injected onto a Halo-C18 column (2.1 × 50 mm, 2.7 μm) at a flow rate of 0.50 mL/min at 50 °C. To remove non-catenane species, solvent, and TFA, a short gradient (60−70% MeOH/H$_2$O containing 0.1% formic acid) was used. A second gradient (70−90% MeOH/H$_2$O containing 0.1% formic acid over 7 min) and a subsequent 90% MeOH/H$_2$O (0.1% formic acid) mobile phase were utilized for an additional 4 min for catenane speciation. See Supporting Information-2 for HPLC traces of DSA solutions and LC-TOF analyses of the generated catenanes.

■ RESULTS

1. [2]-Catenane Generation. Under dynamic self-assembly (DSA) conditions, mixtures of two dipeptide monomers that are epimeric at proline (D-**1a** and L,L-^{H}**2** or ^{H}**2**^{Ph}) self-assemble into octameric [2]-catenanes (Scheme 1)[11]

Scheme 1. [2]-Catenane Self-Assembly from Two Dipeptide Monomers (D-1a and L,L-^{H}2 or ^{H}2^{Ph})
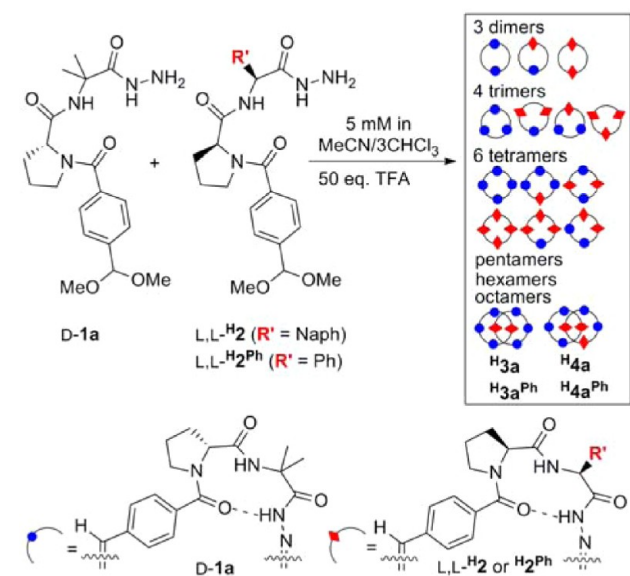
in a diastereoselective and constitutionally selective fashion. X-ray structure analysis[3] showed that the structure contains a dense core from which extend loops terminating in β-turns.

As shown in Figure 1a, the structure is flower-like and assembled from two interlocked saddle-shaped loops where
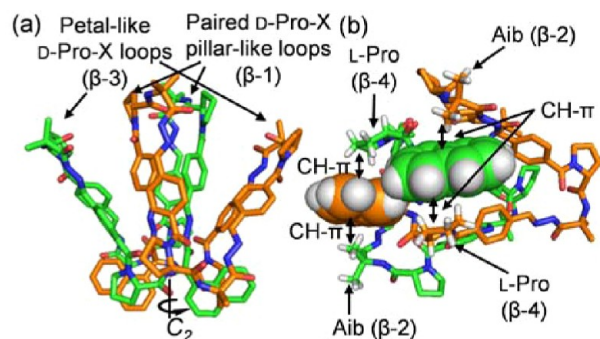
Figure 1. (a) X-ray structure of ^{H}**3a** displaying two identical interlocked tetramers. (b) View from the bottom-side of (a) showing CH−π interactions between the naphthyl (space-filled shape) and L-Pro moiety and between the naphthyl and a methyl group of the Aib residue. All hydrogen atoms except those participating in the CH−π−CH sandwiches are omitted for clarity.

each macrocycle is formed from three units of **1** and one unit of **2**, i.e., [**1**$_3$**2**$_1$]$_2$. X-ray and NMR structural investigation of catenane ^{H}**3a** revealed that its dense self-complementary structure is stabilized by a combination of six intra- and six inter-macrocyclic hydrogen bonds, numerous face-to-face and edge-to-face aromatic interactions, and the previously mentioned CH−π interactions. These interactions were commonly found in all examined catenanes.[3]

Among these non-covalent interactions, the most intriguing were the CH−π interactions which grip the naphthyl ring of naphthylglycine between a proline and an Aib residue (Figure 1b) (Aib = 1-aminoisobutyric acid, or α-methyl-alanine). The rigidity of the peptide loops and the core structure were probed through H/D exchange experiments in protic D-solvents (CD$_3$OD, D$_2$O/pyridine). The exchangeable amide and hydrazone NH groups revealed a variety of chemical environments, including those in the outer loops which exchange readily, and those flanking the core naphthyl, which were much slower. Two key NH's forming inter-ring hydrogen bonds in the core (β-2 and β-4 turns of the catenane ^{H}**3a**; see turn designations in Figure 1b) resisted exchange even after multiple weeks in an aggressive H/D exchange solvent (D$_2$O/pyridine). In other examined catenanes, this H/D exchange eventually occurred but much more slowly than in the outer loops. As fast exchange was taken to reflect structural flexibility and breathability, these observations suggest that the core is rigidified/stabilized by the CH−π interactions, which sandwich the naphthyl rings and pairs of inter-ring hydrogen bonds.

X-ray structure analysis[3] additionally revealed that the core of each catenane is composed of a type II′ β-2 turn and a type VIII β-4 turn. In contrast, the peripheral turns (β-1 and β-3) can adopt either type I′, type I, or type II′ β-turns. The invariability of the core suggests that these β-turn types may also play a role in facilitating the above core interactions.
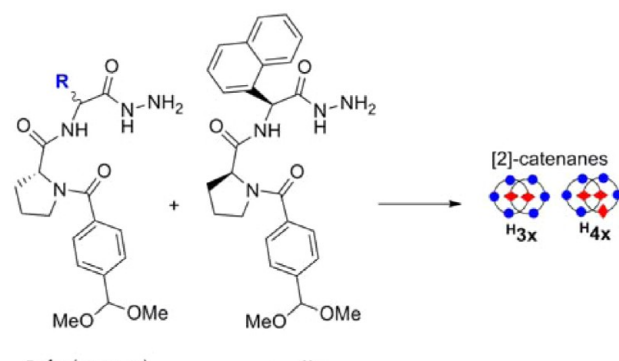
Given the notion that the CH−π interaction was the key differentiating recognition element, we first sought to gain insight by perturbing the β-2 CH donor (Scheme 2b) and

Scheme 2. Illustration of the CH−π−CH Sandwich in ^{H}3a Showing CH−π Interactions (a) between the Naphthyl and L-Pro Moiety (β-4 CH−π) and (b) between the Naphthyl and a Methyl of the Aib Residue (β-2 CH−π)
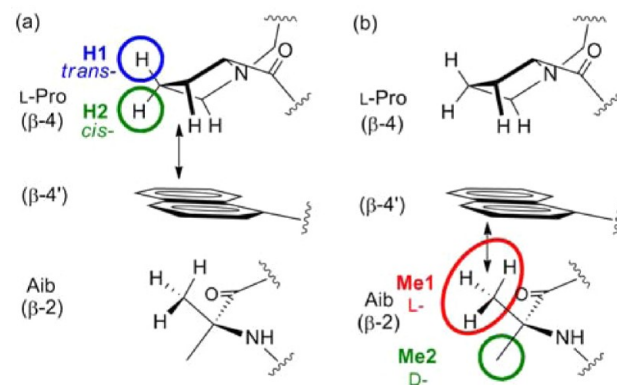
determining how the [2]-catenane responded. To this end, a series of methyl deletion experiments were designed to manipulate the Aib residue of β-2. Since this group originates from the D-Pro-Aib unit, this experiment was accomplished by employing the D-Pro-L-Ala, D-Pro-D-Ala, or D-Pro-Gly monomers for the DSA reaction with monomer L,L-^{H}**2** (Table 1). As described in Scheme 2b, this series of monomers remove the non-interacting methyl group (**Me2**) in the D-Pro-L-Ala case, the interacting methyl group (**Me1**) in the D-Pro-D-Ala case, and both groups in D-Pro-Gly scenario. Not surprisingly, the D-Pro-L-Ala situation (entry 2) gave only a slightly lowered concentration of [2]-catenanes relative to Aib (50 vs 70%), as this methyl deletion combination should leave the CH−π interaction largely untouched. The D-Pro-D-Ala scenario (entry

Table 1. [2]-Catenanes Generation from D-, D,L-, or D,D-1x (x = a, b, v) and L,L-^{H}2

D-1x (x = a, v)
D,L-1b, D,D-1b

L,L-^{H}2

| Entry	R Group	[2]-Catenane yields (%)[a]			
		^{H}3x	^{H}4x	Others[b]	Total
1	a	⌇—	38 (63)	29 (3)	5 (0)	72 (66)
2	L-b	⌇	40 (56)	15 (5)	0 (0)	55 (61)
3	D-b[c]	⌇	0 (0)	0 (0)	0 (0)	0 (0)
4	v	H	31 (50)	7 (4)	1 (0)	39 (54)

[a]Yields obtained from the UV trace of the corresponding catenane in each DSA solution at day 7−9 (steady state) using a calibration curve. The calibration curve was obtained from the HPLC-UV area (289 nm) of several known concentrations of pure catenanes; details in SI1 Figure 1. The values in parentheses were obtained from 3:1 (D:L) biased DSA solutions. [b]Other constitutional isomers of **3x** and **4x**. [c]Only trace amounts of [2]-catenanes were generated (≪ 1%).

3), however, was unexpected, as it led to a large decrease (>1000-fold) in catenanes. While on the surface this suggested that the CH₃−π interactions (there are two C_2-related positions) were stronger than expected from model studies and calculations,[4a,12] this interpretation was immediately questionable, as the D-Pro-Gly case (entry 4), which also lacks the interacting **Me1** group, forms a significant quantity of [2]-catenanes (33%). These data require an explanation that extends beyond the simple CH−π interaction mode.

A likely source for this trend became apparent when the inherent β-turn preference of the core units was considered, as the D-Pro-Aib, D-Pro-D-Ala, D-Pro-L-Ala, and D-Pro-Gly units occupy the invariant β-2 location in the core. Revealing was the fact that D-Pro-L-Ala, D-Pro-Gly, and D-Pro-Aib each prefer type II′ β-turns, while D-Pro-D-Ala prefers a type I′ β-turn.[13] That is, the three monomers preferring a type II′ turn form significant [2]-catenane, while the single case which does not is highly destabilized. The lowered quantity of [2]-catenanes for D-Pro-Gly relative to entries 1 and 2 indicates that, while the CH−π interaction is apparently stabilizing, this effect is subordinate to the ability to adopt a β-2 type II′ turn.

While the stability trends for entries 1, 2, and 4 suggest stability perturbations through CH−π effects, these data stimulated a broader investigation examining β-turn preferences, predisposition to key hydrogen bonds, and CH−π interactions. This study was broken down into two main components, each focusing on a different face of the aryl ring of the CH−π−CH sandwich (β-2 and β-4). Since the five X-ray structures and additional NMR spectra suggest that the core

only accommodates a type II′ β-2 and a type VIII β-4 turn, it become necessary to determine whether the adaptability of the dipeptides to these turn types was a prerequisite for catenane formation. We first address this issue for the β-2 turn in detail using the D-Pro-D-Ala monomer, and then we present the result of additional perturbations to the β-2 face of the aryl ring, simultaneously considering the natural β-turn preferences of the monomer variations (Scheme 3). This modulation sought

Scheme 3. [2]-Catenane Self-Assembly from Pairs of Hydrazine/Acetal Monomers
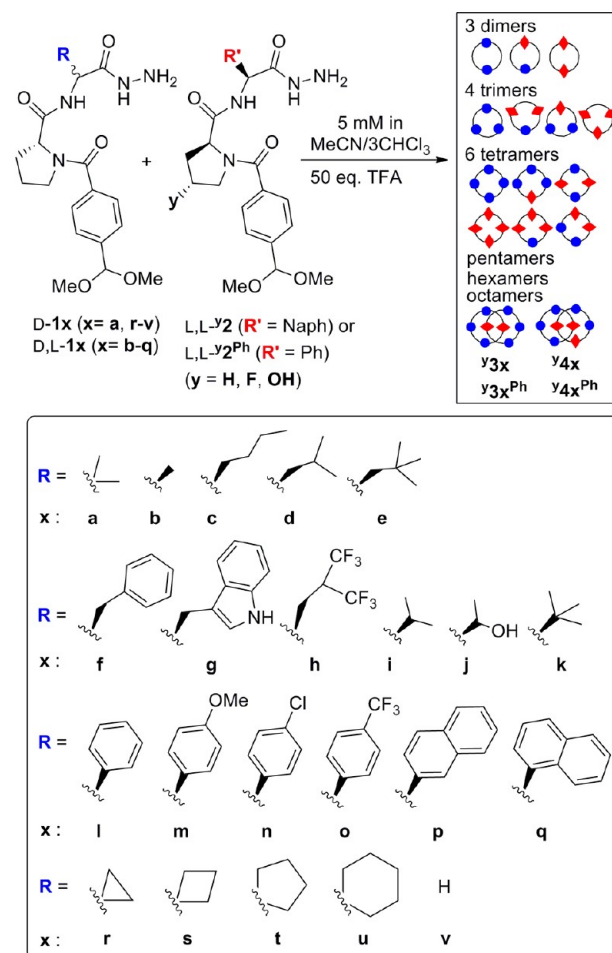
to vary the structure of the β-2 CH donor to determine how catenane stability was affected, vis-à-vis its equilibrium concentration. This was achieved by substituting the D-Pro-Aib unit for a variety of D-Pro-L and D-Pro-cyclic dipeptide structures. Since D-Pro-L-X and D-Pro-cyclo-X dipeptides can adopt type II′ turns,[14] these variants were thought to principally vary the β-2 CH donor of the CH−π−CH sandwich. The β-4 CH donor component of the CH−π−CH sandwich and the turn propensities were also examined through proline variants and the stereoisomeric dipeptides of L-Pro-L-arylGly (^{H}2 and ^{H}2^{Ph}), respectively.

The effect of the above-described mutations on catenane formation/stability was assessed by determining the equilibrium speciation of the DSA solutions. Typical compositions included cyclic dimers through hexamers (pentamers and hexamers were present in low abundance), with each *n*-mer being composed of multiple constitutional isomers of the two monomers. HPLC was used to monitor the approach to the steady state, which varied for different L-monomers.[15] In most cases octameric [2]-catenanes are visible within 5 min and a significant amount of catenane accumulates within a day (SI1 Figure 2). Notably, the tetrameric macrocycle corresponding to one half of **3x** (i.e., $[1_32_1]$) is barely detectable, and one of the most abundant cyclic species at early time, a trimer subunit of **3x** composed of two units of **1** and one unit of **2** ($[1_22_1]$), substantially decreases as the cantenane concentrations increase.

To compare [2]-catenane yields across the series, [2]-catenane concentrations were determined from the HPLC−UV trace (289 nm) and calibration curves established for several purified [2]-catenanes; compound-to-compound differences in extinction coefficient were negligible (SI1 Figure 1). In several instances a precipitate was noted at the 3−5 day time point of the DSA reaction. Even when this occurred, the catenane concentration in the supernatant either continuously increased to a steady state, which then began to slowly drop after 15 days, or reached a steady state before the onset of the precipitation. These non-equilibrium situations are noted.

By quantifying the amount of catenane in a DSA reaction, one measures how its free energy compares to that of the other components of the mixture. In addition to affecting the catenane's energy, varying the DSA components could also perturb the relative energy of the other components of the equilibrium mixture (dimers, trimers, etc.), and thus hinder the comparability of one system to another. To assess the viability of cross-system comparisons, we carried out full speciation analyses on several DSA reactions (D-**1a**/L,L-**H2**, D,L-**1b**/L,L-**H2**, D,D-**1b**/L,L-**H2**, D-**1v**/L,L-**H2**, D-**1a**/L,L-**H2Ph**, D,D-**1b**/L,L-**H2Ph**, and D-**1v**/L,L-**H2Ph** in Scheme 3) and looked for any unusually stable (or unstable) cyclic oligomers that could artificially lower (or raise) the amount of catenanes (see representative cases in SI1 Figure 3−5). These experiments demonstrate that the dimers:trimers:tetramers ratio does not change significantly from system to system, and suggest that the structural changes being explored herein do not significantly affect the relative energy of the non-catenated, less structured, cyclic oligomers present in the DSA solution. With the caveat that this analysis does not rigorously rule out subtle effects, we report both absolute and relative (to the other DSA components) concentrations of the [2]-catenanes and interpret these data to reflect [2]-catenane stability.

2. β-2 Turn and [2]-Catenane Generation. *2.1. Turn Type.* X-ray structure analysis revealed that all examined catenanes adopt a type II′ β-2 turn in the core, obviously suggesting that this turn is most suited to a stable [2]-catenane. The combination of a type II′ β-2 and a type VIII β-4 creates a core structure wherein two key interactions can take place: (1) CH−π interactions between the X residue of the D-Pro-X unit and the aryl residue of the L-Pro-L-arylGly unit and (2) two inter-ring H-bonds between two β-turn backbones (see red-boxed interactions in Figure 2). As mentioned previously, the preferred β-turn of the D-Pro-D-Ala dipeptide is type I′. Superimposing this mismatched turn geometry onto the structure of **H3a** reveals several destabilizing interactions. In the type I′ turn the amide is rotated nearly 180°,[16] which
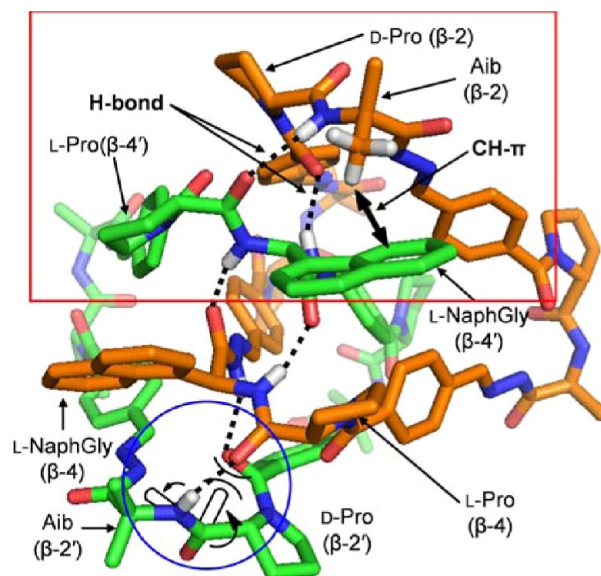
Figure 2. X-ray structure of **H3a** displaying the inter-ring hydrogen bonds (black dot-lines). The red box indicates a CH−π interaction (black up-down arrow) and inter-ring hydrogen bonds between the β-2 and β-4 backbones. The blue circle indicates the loss of the inter-ring hydrogen bond and gain of the repulsive interaction caused by a proximate carbonyl group in the β-2 turn when a type I′ turn instead of a type II′ turn is accommodated in the β-2 turn. All hydrogen atoms except those forming inter-ring hydrogen bonds and the CH−π interaction (red box) are omitted for clarity.

removes the inter-ring hydrogen bond while simultaneously creating a repulsive carbonyl−carbonyl interaction (see the shadow interactions in Figure 2). This repulsive term, combined with the loss of two CH₃−π interactions (doubled due to C_2 symmetry), is the likely cause for the dramatic loss of [2]-catenane (>1000-fold drop).[17]

Interestingly, a similar degree of destabilization can be found in the difference in binding affinity of vancomycin to the D-Ala-D-Lac mutant relative to D-Ala-D-Ala (1000-fold drop).[18] The introduction of an ester oxygen in D-Ala-D-Lac causes both the loss of a hydrogen bond and the introduction of a destabilizing electrostatic repulsion between a carbonyl and the ester oxygen in D-Ala-D-Lac. This parallel suggests that the loss of the hydrogen bond and the additional repulsive interaction expected for D-Pro-D-Ala may be the primary reasons for destabilization, with loss of the CH₃−π interactions playing a lesser role. This is supported by the ability of the catenane to accommodate D-Pro-Gly in the core.

Although the above structural rationale for loss of catenane using D-Pro-D-Ala is strong, two additional explanations were considered: (1) an exceptionally stable cyclic oligomer was being formed, with the concomitant effect of draining the concentration of [2]-catenane, and (2) the loss of stability was not due to a specific set of interactions at the β-2 position, but the monomer simply could not occupy one of the other two positions (β-1 and β-3) in the structure (despite their heightened structurally flexibility).

To address scenario 1, a full speciation analysis was carried out on the non-catenated products for the two experimental permutations: D-Pro-Aib/L-Pro-L-NaphGly (D-**1a**/L,L-**H2**), which gives [2]-catenane, and D-Pro-D-Ala/L-Pro-L-NaphGly (D,D-**1b**/L,L-**H2**), which does not. The results clearly demonstrate that no unusually stable or unstable *n*-mers exist in either

experiment (see SI1 Figures 3 and 4). When catenane is formed, this naturally lessens the total amount of n-mers remaining at equilibrium, but within this group of cyclics, the distribution of 2-mers:3-mers:4-mers is unchanged.[19] Since no unusual stabilities were evident in the macrocycles, these data suggest that the lack of [2]-catenane formation with D-Ala lies in the [2]-catenane itself.

Given this evidence, the problem logically reduces to the D-Ala monomer not being accommodated into the β-1, β-2, and β-3 positions or combinations thereof. The turn type destabilization rationale presented above would suggest that the problem is localized at the β-2 site, but since the data do not rule out the possibility that the problem lies with its positioning at β-1, β-3, or both, three monomer experiments were designed. Outlined in Scheme 4 is the logical consequence of mixing monomer **2** with two different **1** monomers on the production of **3**,[20] one of which is the previously unincorporated D-Pro-D-Ala (D,D-**1b**), while the second is the successfully incorporated D-Pro-Aib (D-**1a**).

In the event that D-Pro-D-Ala is not incorporated in any of the three positions (i.e., NO, NO; case 1 in Scheme 4), then the three monomer experiment would form [2]-catenane incorporating the D-Pro-Aib monomer exclusively, i.e., [(D-**1a**)$_6$(L,L-^{H}**2**)$_2$], with no catenanes containing D,D-**1b** being detected. If, on the other hand, the D-Pro-D-Ala monomer is incorporated into one or a combination of the β-sites, then different constitutional isomers are possible, each of which can be discriminated in an LC-TOF analysis. While the β-1 and β-3 sites are chemically inequivalent, the analysis in Scheme 4 treats them as either equally probable or improbable, though their contributions to the constitutional isomerism of mixed [2]-catenanes is implied. As will be seen below, the logical differentiation between the β-1 and β-3 sites was not necessary to explain the data.

If the D-Pro-D-Ala monomer could only occupy the β-2 sites (i.e., YES, NO; case 3 in Scheme 4), only two *mixed* catenanes of **3** would be formed. If the opposite were true and D-Pro-D-Ala was only *excluded* from the β-2 turn (i.e., NO, YES; case 2 in Scheme 4), then a significant number of mixed [2]-catenanes would result. For [2]-catenane **3**, this latter scenario would require that each isomer contain a *minimum* of two D-Pro-Aib monomers to occupy the β-2/2′ positions, with the balance of positions being occupied by logical combinations of D-Pro-Aib and D-Pro-D-Ala. Conversely, a *maximum* of four D-Pro-D-Ala units can be accommodated (case 2 in Scheme 4).[21]

When D-Pro-D-Ala, D-Pro-Aib, and L,L-^{H}**2** were combined in a 1:1:2 ratio, a complex mixture of catenanes was obtained, though the identity of each compound could be deconvoluted by LC-TOF (Figure 3). Clear from this analysis was the fact that each catenane contained at least 2 equiv of L,L-^{H}**2** (expected) and *no less than 2 equiv* of D-Pro-Aib. The diversity in composition in Figure 3a, therefore, results from the D-Ala and Aib monomers being statistically distributed into the β-1, β-1′, β-3, and β-3′ sites, with the β-2 and β-2′ sites being solely occupied by the Aib monomers.
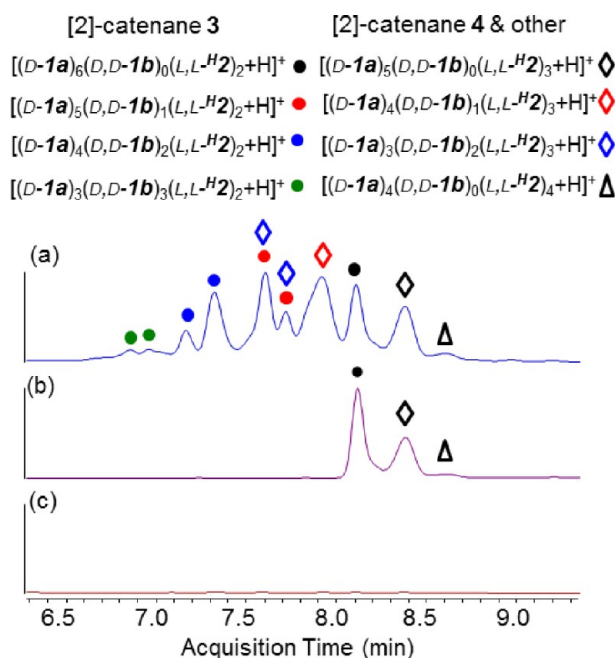
When the ratio of D-Pro and the L-Pro monomers is changed to 3:1, that is, D-Pro-D-Ala:D-Pro-Aib:L,L-^{H}**2** = 3:3:2, the catenane compositional diversity is unchanged from Figure 3a, though the absolute intensity of each species increases due to the D-Pro-bias (SI1 Figure 6a). On the other hand, varying the ratio of the two D-Pro monomers (keeping the D-Pro:L-Pro ratio at 1:1) shows that D-Pro-D-Ala-rich cases generate less catenanes and fewer isomers, while D-Pro-Aib-rich cases

Scheme 4. Logical [2]-Catenane Combinations from Mixing Two 1 Monomers (D-1a, D,D-1b) with the L,L-^{H}2 Monomer
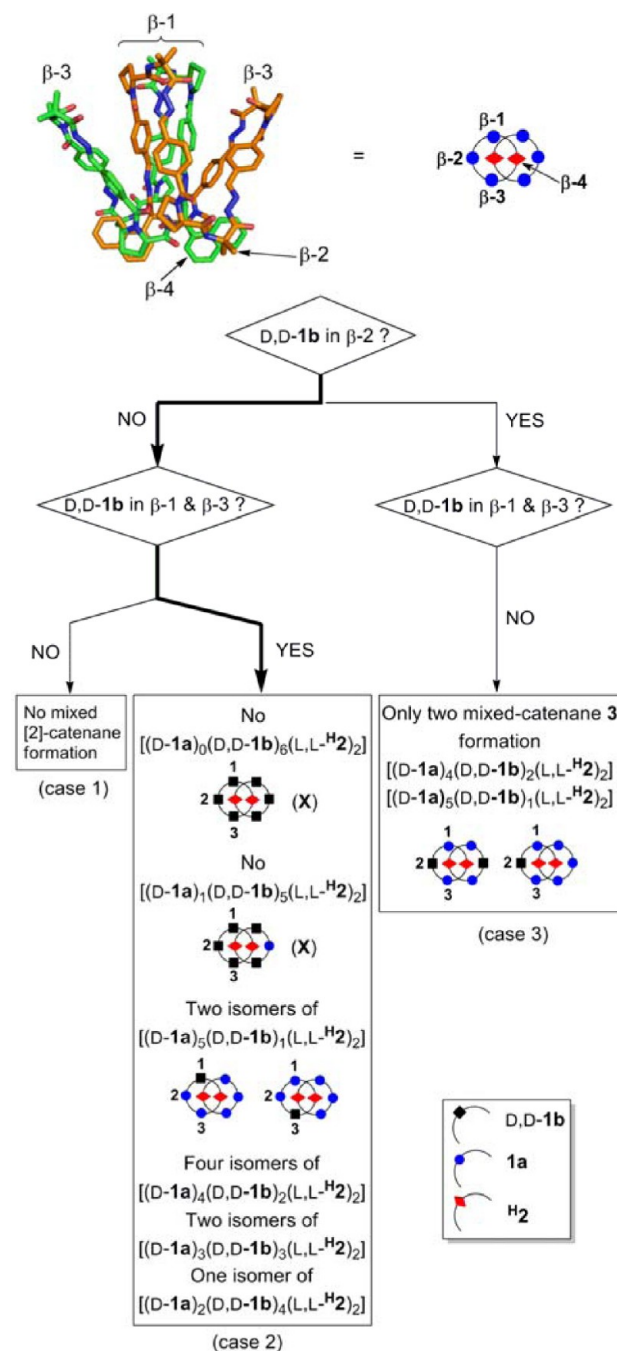
generate comparatively more catenane but fewer isomers (SI1 Figure 6b).

We interpret these data to reflect an absolute requirement for two L,L-^{H}**2** monomers in the β-4/4′ positions and two D-Pro-Aib monomers in the β-2/2′ positions.[22] The D-Pro-D-Ala monomer can therefore be easily accommodated into the β-1 and β-3 sites, but *not* the β-2 site. We conclude that the loss of catenane stability is localized to the β-2/2′ sites, and that the principal requirement for catenane stability is adopting the necessary type II′ β-turn. Once the turn is in place, the noncovalent interactions in Figure 2 function to perturb the comparative stabilities.

[2]-catenane **3**

$[(\text{D-}\mathbf{1a})_6(\text{D,D-}\mathbf{1b})_0(\text{L,L-}^H\mathbf{2})_2+\text{H}]^+$ ●

$[(\text{D-}\mathbf{1a})_5(\text{D,D-}\mathbf{1b})_1(\text{L,L-}^H\mathbf{2})_2+\text{H}]^+$ ● (red)

$[(\text{D-}\mathbf{1a})_4(\text{D,D-}\mathbf{1b})_2(\text{L,L-}^H\mathbf{2})_2+\text{H}]^+$ ● (blue)

$[(\text{D-}\mathbf{1a})_3(\text{D,D-}\mathbf{1b})_3(\text{L,L-}^H\mathbf{2})_2+\text{H}]^+$ ● (green)

[2]-catenane **4** & other

$[(\text{D-}\mathbf{1a})_5(\text{D,D-}\mathbf{1b})_0(\text{L,L-}^H\mathbf{2})_3+\text{H}]^+$ ◇

$[(\text{D-}\mathbf{1a})_4(\text{D,D-}\mathbf{1b})_1(\text{L,L-}^H\mathbf{2})_3+\text{H}]^+$ ◇ (red)

$[(\text{D-}\mathbf{1a})_3(\text{D,D-}\mathbf{1b})_2(\text{L,L-}^H\mathbf{2})_3+\text{H}]^+$ ◇ (blue)

$[(\text{D-}\mathbf{1a})_4(\text{D,D-}\mathbf{1b})_0(\text{L,L-}^H\mathbf{2})_4+\text{H}]^+$ △
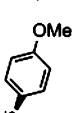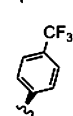
Figure 3. LC-TOF UV trace (289 nm) of [2]-catenanes at day 7 generated from (a) D-**1a**:D,D-**1b**:L,L-^{H}**2** = 1:1:2 mixture, (b) D-**1a**:L,L-^{H}**2** = 1:1 mixture, and (c) D,D-**1b**:L,L-^{H}**2** = 1:1 mixture (5 mM in 25% MeCN/CHCl$_3$, 50 equiv of TFA). Each peak was identified by LC-TOF analysis.

2.2. Aib Modifications in D-Pro-Aib. The X-ray structure analysis and ROESY NMR spectrum of ^{H}**3a**[3,11] established a CH−π interaction between the nucleated naphthyl group and one of the Aib methyl groups (β-2). Since only one CH$_3$ group was upfield shifted, it implied that [2]-catenane stability might not be restricted to Aib, as other arrangements should be capable of creating a productive CH−π interaction, provided that the preferred β-turn is type II′ (vide supra). The catenanes generated by the D-Pro-L-Ala monomer (entry 2 in Table 1) support this hypothesis. Candidate structures for varying the β-2 CH donor include D-Pro-X and D-Pro-L-X dipeptides, as these predominantly form type II′ β-turns.[13c,23]

As shown in Table 2 and SI1 Table 1, all tested D- and D,L-monomers generated the corresponding [2]-catenanes, with both aliphatic and aromatic groups being well behaved. In the D,L-monomers having an aliphatic alanine substituent, a clear steric trend was noted, with larger groups destabilizing the structure. Considering the dense nature of this region (Figure 4), the stability response to steric changes (n-Bu vs i-Pr vs tert-Bu) was surprisingly muted, as even a tert-butyl for methyl substitution provided a measurable quantity of [2]-catenane (Table 2, entry 7). Polar substituents (OH or F), however, were disproportionally detrimental to stability.[24] This observation is readily accommodated by a model which invokes a repulsive polar−π interaction in the constrained pocket.

D,L-Monomers having an arylGly residue also generated the corresponding [2]-catenanes (Table 2, entries 8−10; SI1 Table 1, entries 4−6), implying that a stabilizing π−π interaction between the naphthyl group and the arylGly residue can be accommodated. One example of this family of [2]-catenanes was crystallographically characterized[3] and shown to adopt a nonsymmetric solid-state structure featuring a CH−π−π sandwich in the core, wherein the β-2 face creates π−π interactions that are either of the slipped π−π or edge-to-face

Table 2. Yields of [2]-Catenanes in the DSA Solutions Formed from Various D-Monomers and L,L-^{H}2

Entry	R Group	[2]-Catenane yields (%)[a]			
		^{H}**3x**	^{H}**4x**	Others[b]	Total
1	**c**[c]		39 (-)	25 (-)	3 (-)	67 (-)
2	**d**		35 (60)	28 (7)	3 (0)	66 (67)
3	**g**		63 (24)	0 (0)	0 (0)	63 (24)
4	**h**[c]		10 (-)	5 (-)	0 (-)	15 (-)
5	**i**[c]		39 (-)	11 (-)	0 (-)	50 (-)
6	**j**		29 (15)	0 (0)	0 (0)	29 (15)
7	**k**		8 (14)	3 (3)	0 (0)	11 (17)
8	**l**		42 (54)	11 (2)	3 (3)	56 (59)
9	**m**		42 (55)	14 (4)	2 (4)	58 (63)
10	**o**[c]		24[d] (-)	6[d] (-)	1[d] (-)	31[d] (-)
11	**r**		16 (23)	20 (9)	2 (0)	38 (32)
12	**s**		22 (56)	25 (2)	22 (4)	69 (62)
13	**t**		26 (50)	31 (10)	18 (6)	75 (66)
14	**u**		28 (54)	30 (10)	17 (4)	75 (68)

[a]Yields obtained from the UV trace (289 nm) of the corresponding catenane in each DSA solution at day 7−9 using a calibration curve. The calibration curve was obtained from the HPLC-UV area (289 nm) of several known concentrations of pure catenanes; Details in SI1 Figure 1. The values in parentheses were obtained from 3:1 (D:L) biased DSA solutions. [b]Other constitutional isomers of **3x** and **4x**. [c]Biased DSAs (3:1) led to precipitates and yields are not reported. [d]Precipitation occurred within 3−5 days. The yields of the generated catenanes were obtained at day 7−8 after filtering the precipitates.

variety. In cases where electronically perturbed phenyl glycines were utilized, the electron-releasing *para*-methoxy substituent gave more catenane than did the p-CF$_3$ substituent (Table 2, entries 9 vs 10), in apparent contradiction to the Hunter−Sanders model.[25]

When cyclic α,α-dialkylated analogues of Aib were used, outcomes that depended on ring size were obtained, with the optimal size being a C$_5$ or C$_6$ ring (Table 2, entries 11−14). As predicted by the respective catenane stabilities, the β-turns of D-
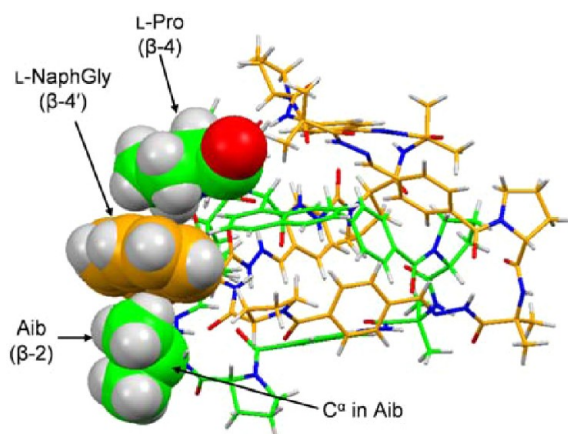
Figure 4. X-ray structure of ^{H}**3a** displaying the compact structure between the naphthyl and the Aib. Space-filled shapes are used for these groups.

Pro-1-amino-1-cycloalkanecarboxylic acids (Ac$_5$c and Ac$_6$c) are similar to those of D-Pro-Aib.[14d,26] Although the reasons are not obvious, these cyclic monomers skew the catenane compositions toward ^{H}**4x** and the other minor catenanes. Since the cyclic α,α-dialkylated analogues of Aib populate a more constrained conformational space at both the mutated and adjacent positions,[14e,27] these modifications apparently compensate for the structural changes that result from replacing a type II' for a type I turn at the β-3 or β-1 positions.

As mentioned in section 1, the D-Pro-Gly monomer, which lacks the methyl group that participates in the CH−π interaction but favors a type II' β-turn, creates a significant quantity of [2]-catenane under the DSA conditions.[28] The structural flexibility noted by the steric trend in this position suggests that the structure could adjust to fill the void and gain stabilization through a backbone CH−π interaction. Alternatively, the competing cyclic n-mers would somehow have been disadvantaged, though a speciation analysis suggests that this is not the case (SI1 Figure 5). A three-component DSA reaction using a 1:1 mixture of D-Pro-Gly/D-Pro-Aib as the **1** components with L,L-^{H}**2** returned results indicating that, like D-Pro-Aib, D-Pro-Gly has an equal propensity to occupy the β-1, β-2, and β-3 positions (in contrast to D-Pro-D-Ala; see SI1 Figures 8−10).[29] In all respects the D-Pro-Gly monomer is statistically equivalent to the parent monomer (D-Pro-Aib) and is thermodynamically viable in all positions. These data therefore imply that a CH−π interaction on the β-2 face of the arene is not an absolute requirement for competitive stability or the structure can significantly reorganize to position the glycine CH$_2$ group into the CH$_3$−π position.

In an attempt to separate the type II' β-2 turn requirement from the subsequent inter-ring hydrogen bond, the D-Pro-L-Pro monomer was tested. This dipeptide favors a type II' β-turn[30] but lacks the amide NH to complete the inter-ring hydrogen bond (Figure 5). The ~10% yield of catenanes (see SI1 Table 1) indicates that a structure lacking two (β-2(NH)→β-4(C═O)) of six inter-ring hydrogen bonds can at least partially compete with non-catenated macrocycles, though at a price. The generation of catenanes with this disadvantage implies that the otherwise conserved inter-ring hydrogen bond is not an absolute requirement for catenane stability, as its loss can be partially offset by the network of CH−π, π−π, and inter-ring hydrogen-bonding interactions.
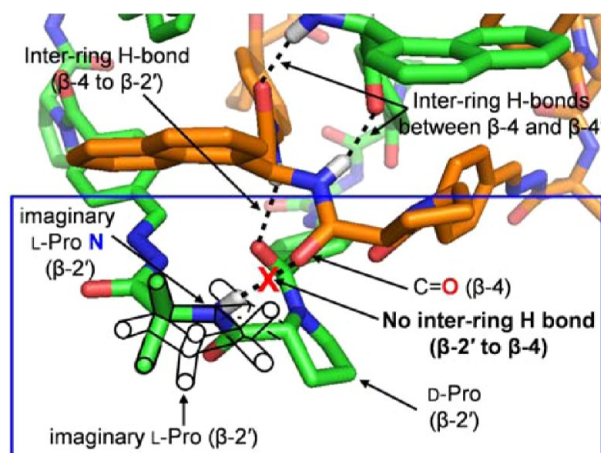
Figure 5. Picture superimposing an imaginary L-Pro group instead of the Aib (β-2') in the X-ray structure of ^{H}**3a** (the black outlined shape in the blue box). Since L-Pro group has the N (β-2') instead of the amide NH, it induces a structure lacking an inter-ring hydrogen bond (β-2'(NH)→β-4(C═O)) between the β-2 and β-4 backbone. All hydrogen atoms except those forming inter-ring hydrogen bonds are omitted for clarity.

3. β-4 Turn and [2]-Catenane Generation. 3.1. Turn Type. X-ray structure analysis and H/D exchange experiments[3] indicated that every β-4 turn in the core is of the type VIII form and this conformation is maintained in solution. Previous studies have shown that L-Pro-L-X dipeptide units favor type I turns,[31] which is the form adopted by L-Pro-L-NaphGly when it is located in the flexible β-3 loop of catenane ^{F}**4u** (^{F}**4b** in ref 3). In this position, the naphthyl group is not involved in any CH−π interactions.

Adopting a type VIII turn in β-4 thus requires the L-Pro-L-arylGly units to undergo a turn change. A comparative analysis of the standard turn dihedral angles (type I, Φ_{i+1}, −64°, Ψ_{i+1}, −27°, Φ_{i+2}, −90°, and Ψ_{i+2}, −7°; type VIII, Φ_{i+1}, −72°, Ψ_{i+1}, −33°, Φ_{i+2}, −123°, and Ψ_{i+2}, 121°) indicates that the conversion from type I to type VIII requires an ~130° rotation of the $i+2$ carbonyl group (Ψ_{i+2}), which links the L-arylGly carbonyl group to the 1,4-benzamide of the following turn. This analysis suggests that the energy difference between type I and type VIII is sufficiently small as to be compensated by the CH−π interactions between the Pro and arylGly residues and two inter-ring hydrogen bonds (β-4(N)→β-4'(O) and β-4'(N)→β-4(O), Figure 6).

Stereoisomeric L-Pro-L-arylGly (aryl = naphthyl, ^{H}**2**; phenyl, ^{H}**2^{Ph}**) dipeptides, i.e., L-Pro-D-arylGly, D-Pro-D-arylGly, and D-Pro-L-arylGly dipeptides, were investigated to probe the validity of the above suggestion. In combination with D-Pro-Aib (D-**1a**), each of these cases generated only non-catenated cyclic macrocycles. Since the $i+1$ and $i+2$ configurations of the L,D, D,D, and D,L dipeptides favor type II, I', and II' β-turns, respectively,[13c,23,30c] their metrical biases[32] would require significant conformational change to both residues (Pro and arylGly) to achieve the optimal β-4 structure, differences which might not be compensated for through weak secondary forces.

3.2. L-Pro Modification in the L-Pro-L-arylGly unit. As depicted in Figure 1 and Scheme 2a, the structure of ^{H}**3a** supports a CH−π interaction between a proline unit and the naphthyl group (β-4 to β-4'). Since the acidity and geometric arrangement of these CH donors should perturb the CH−π interaction,[33] we investigated the effect of electronegative
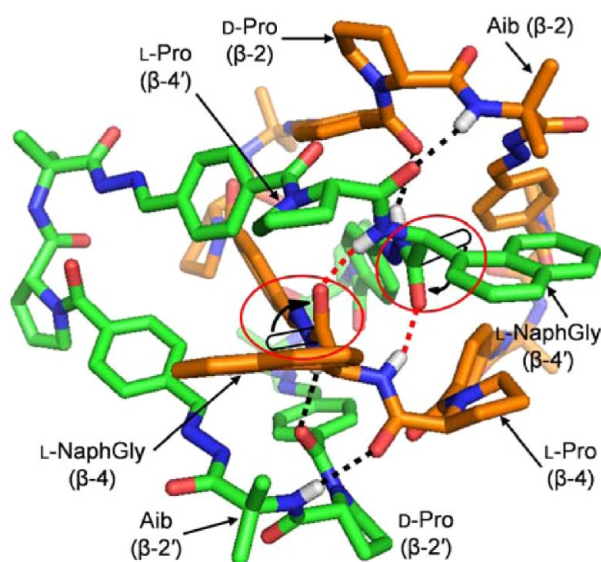
Figure 6. X-ray structure of ^{H}**3a** displaying the inter-ring hydrogen bonds between the β-4 and β-4′ backbones (red dotted lines). The red circles indicate the rotation (the black arrow) of the L-NaphGly carbonyl group (i+2 residue) in the β-4 turn for the turn conversion (type I to type VIII). All hydrogen atoms except those forming inter-ring hydrogen bonds are omitted for clarity.

substituents at **H1** and **H2** (Scheme 2a). Supporting this hypothesis are experimental studies by Zondlo showing that inductive effects enhance proline CH−π interactions,[12a] and computational studies by noting the benefit of α-halo substituents on CH−π binding strengths.[33b]

To this end we examined DSA reactions with mixtures of D-Pro-Aib (D-**1a**) and variants of L-^{y}Pro-L-NaphGly (y = OH, F), where **H1** is either OH or F (*trans*-y-L-proline; (2S,4R)-4-y-L-proline) or where **H2** is OH (*cis*-hydroxy-L-proline; (2S,4S)-4-hydroxy-L-proline) (Scheme 2a), or similarly, using the L-^{y}Pro-L-PhGly analogues.[34] In all cases, DSA reactions with polar **H2** substituents (i.e., *cis*-OH or *cis*-F) provided no detectable catenane. However, in the **H1** replacement series (*trans*-OH and *trans*-F), significant quantities of the [2]-catenanes were generated (Figure 7).[35] Complicating this analysis was the onset of precipitation after 3−5 days for these monomers, though the kinetic evolution of catenane formation clearly shows that [2]-catenane concentrations reach a steady state before the onset of precipitation (see SI1 Figures 13 and 14).

Since electronegative groups at **H1** (*trans*-y) increase the positive charge on the interacting hydrogens, this substitution should intensify the electrostatic component of the CH−π interaction.[33a,36] An additional factor that may play a minor role is the fact that *trans*-4-substituted prolines favor the C$_\gamma$-exo ring pucker, which is the observed pucker in all structurally characterized catenanes.[3,37] The growth of [2]-catenane concentration on introducing electronegative elements at **H1** implies a slightly stronger CH−π interaction. The perturbation is especially noticeable when the nucleating aromatic is phenyl. The lack of catenane with electronegative **H2** substituents (*cis*-y) likely reflects a destabilizing electrostatic repulsion between the naphthyl π system and the proximal OH or F atom. The C$_\gamma$-endo arrangement of *cis*-4-substituted prolines[37] likely results in an additional destabilization both by depopulating the more favorable pucker, and by also positioning the electronegative group closer to the aryl π−donor. However, since the proline
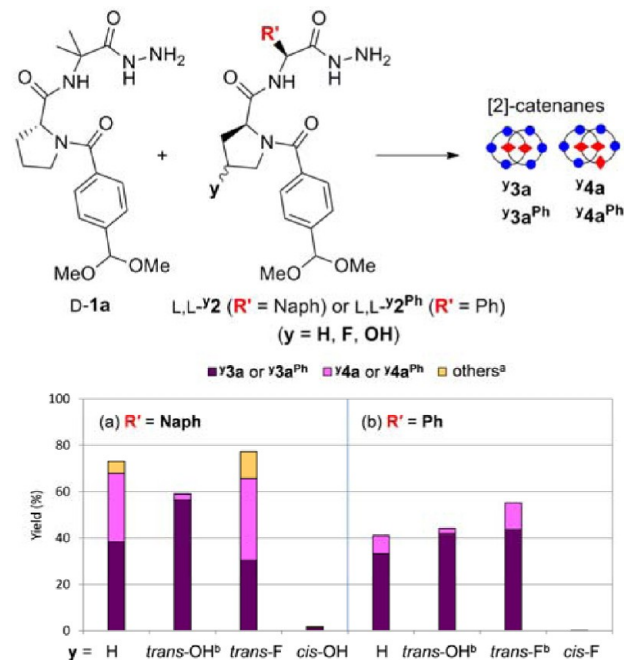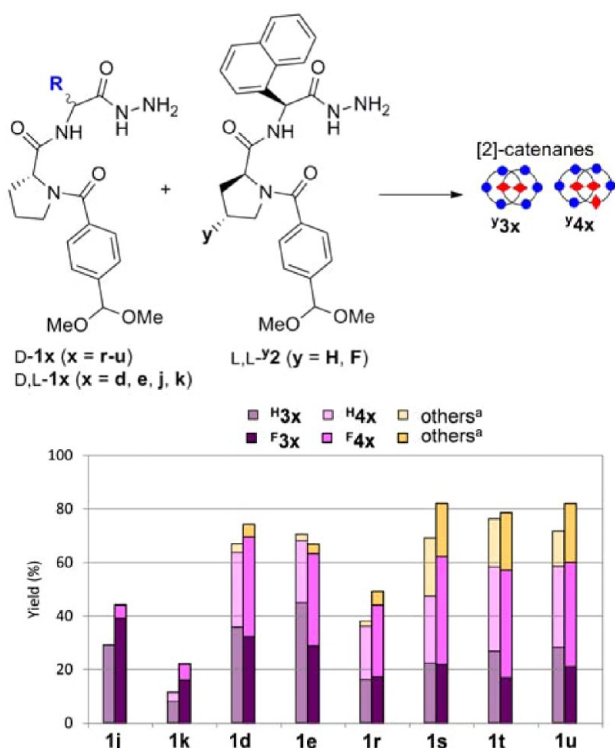

Figure 7. Generation of [2]-catenanes from D-**1a** and L,L-^{y}**2** or L,L-^{y}**2**^{Ph}. ^{a}Other constitutional isomers of **3x** and **4x**. ^{b}Although precipitation of noncatenated products occur after 3−5 days, the catenane reaches a steady state within 3 days (see the details in SI1 Figures 13 and 14).

ring pucker has been known to bias the backbone torsional angles of adjacent residues,[38] it cannot be excluded that *cis*-4-substituted prolines with the C$_\gamma$-endo arrangement prefer a β-turn that is not type I, thus causing a destabilization of the catenane (vide supra).[39]

Unexpectedly, the constitutional selectivity of the catenane is sensitive to the nature of the **H1** substituent, with *trans*-OH decreasing the amount of **4** and *trans*-F increasing it. In the former case, ^{HO}**3a** accounts for more than 95% of the total catenanes. The higher propensity for forming the non-C_2 product **4** in the *trans*-F case was consistently observed in DSA reactions of D-monomers and L-$^{trans\text{-}F}$Pro-L-NaphGly (L,L-$^{trans\text{-}F}$**2**, Figure 8). In all paired DSA reactions (Figure 8), the equilibrium concentration of the non-symmetric [2]-catenane, [^{F}**4x**]$_{eq}$, was higher than [^{H}**4x**]$_{eq}$; the total yield of ^{F}[2]-catenanes was also higher. In particular, ^{F}**4x** accounted for ~50% of all catenanes in DSA reactions generated from cycloaliphatic monomers (D-**1r**−D-**1u**, Scheme 3). The intrinsic preference for ^{H}**4x** in these monomers (Table 2, entries 11−14) and ^{F}**4x** favoritism in fluorine-substituted monomers is additive, as judged by the high [2]-catenane concentrations for **1s**, **1t**, and **1u** in Figure 8. This effect was utilized to advantage in isolating a pure sample of ^{F}**4u** for X-ray analysis.[3]

The X-ray structures in the previous paper[3] indicate that the proline−π interaction is created from two pseudo-axial CH's and one pseudo-equatorial CH interacting with the aryl ring. Computational studies of CH−π interactions between cyclic hydrocarbons and aromatics indicate that CH−π strengths track with the number of contacting CH's, although the total strength is not precisely additive.[40,41] To probe the effect of CH donor number, the four-membered analogue of L-Pro was examined. L-Azetidine-2-carboxylic acid (Aze) was employed (L,L-^{H}**2**$_{Aze}$ and L,L-^{H}**2**$^{Ph}_{Aze}$) along with D- or D,L-**1x** (x = a, b, m, u, v) to determine if reducing the number of CH donors would decrease the amount of [2]-catenanes. This substitution is not

Figure 8. Generation of [2]-catenanes from D- or D,L-**1a** and L,L-^{H}**2** (dot-filled) or L,L-$^{trans-F}$**2** (filled). ^{a}Other constitutional isomers of **3x** and **4x**.

Table 3. [2]-Catenanes Generation from D- or D,L- (x = a, b, m, u, v) and L,L-^{H}**2**$_{Aze}$

Entry	R Group	[2]-Catenane yields (%)^a			
		^{H}**3x**$_{Aze}$	^{H}**4x**$_{Aze}$	Total	$\Delta_{Pro-Aze}$^{b}
1	a	45 (51)	2 (2)	47(53)	25 (13)
2	b	29 (31)	0 (0)	29 (31)	26 (30)
3	m	5^d (21)	0^d(0)	5^d (21)	53 (42)
4	u	55 (67)	4 (1)	62 (69)^c	13 (1)
5	v	H	2 (5)	0 (0)	2 (5)	37 (49)

^{a}Yields obtained from the UV trace of the corresponding catenane in each DSA solution at day 7−9 using a calibration curve. The calibration curve was obtained from the HPLC-UV area (289 nm) of several known concentrations of pure catenanes; details in SI1 Figure 1. The values in the parentheses were obtained from 3:1 (D:L) biased DSA solutions. ^{b}The yield difference between total catenanes obtained from DSAs employing L,L-^{H}**2** and L,L-^{H}**2**$_{Aze}$, respectively. ^{c}The yields of other constitutional isomers of **3x** and **4x** were 3 (1)%. ^{d}Precipitation occurred within 3−5 days. The yields of the generated catenanes were obtained at day 7−8 after filtering the precipitates.

flawless, as Aze's dihedral preferences and the projection of the CH's could be slightly different due to its flattened ring pucker.[42] In addition, a different turn preferences from L-Pro-X cannot be excluded, as the turn preferences of the dipeptide L-Aze-X are not known.

L,L-^{H}**2**$_{Aze}$ and several monomers (Table 3) were combined and the total steady-state concentrations determined. In general the Aze monomer was tolerated, implying that L,L-^{H}**2**$_{Aze}$ can adopt the type VIII turn that L,L-^{H}**2** does. In the parent D-Pro-Aib case, the decreased quantity of [2]-catenanes (from 72% to 47% of the total) implies a weakened CH−π interaction due to either an imperfect arrangement or a reduction of the number of CH donors. For entries 3 and 5, which position a 4-MeO-C₆H₄-ring and a glycine on the opposite face of the naphthyl ring, only traces of [2]-catenane were observed. When the naphthyl group was replaced with phenyl (L,L-^{H}**2**^{Ph}$_{Aze}$), all cases gave <11% [2]-catenanes (SI1 Table 3). Thus, weakening the π, reducing the number of proline CH donors, and removing the optimal CH donors in the β-2 position (Scheme 2) causes total or near total loss of [2]-catenane viability. For the naphthyl cases, it appears that a minimum of three CH−π interactions are needed for measurable stability; a Pro−π−Gly case can be accommodated, but an Aze−π−Gly cannot. Increasing the CH donors from the β-2 position in Aze−π−Aib restores measurable stability, suggesting that the catenane can compensate for the loss of some CH−π interactions, but once a threshold of three has been reached, the structure is no longer competitive. When the π is phenyl (**2**^{Ph} series), the threshold is higher.

■ DISCUSSION

In summary, mixtures of two dipeptides that are epimeric at proline generate octameric [2]-catenanes under dynamic self-assembly (DSA) conditions; the reactions are diastereoselective and constitutionally selective. The solid-state and NMR structural analysis of the [2]-catenanes discussed in the previous paper[3] reveal that the catenanes are stabilized by a combination of intra- and inter-macrocyclic hydrogen bonds, multiple aromatic interactions, and CH−π interactions. The outcome of our structure/stability studies indicates that the core structure is especially important, with the pair of CH−π−CH sandwiches, the inter-macrocycle hydrogen bonds between the β-2 and β-4 turns, and the predisposition for type II' (for β-2) and type VIII (for β-4) β-turns being important variables. Indeed, the β-turn conformation at the core was found to be the critical feature predicting stability, while the others individually contribute to catenane stability; to various degrees they could be weakened or removed with a concomitant cost.

Mutation studies in the core found that, when the proper turns were accessible, the stability of the catenane relative to its competing non-catenated structures could be reasonably

described by a combination of non-covalent interactions. Through methyl deletion experiments, proline to F-proline or HO-proline mutations, and proline to azetidine carboxylic acid mutations, it became apparent that a minimum of three CH−π interactions (doubled by symmetry) are needed to compete with non-catenated structures. In this regard, the CH−π interactions in the pair of core CH−π−CH sandwiches significantly contribute to product stability. Sufficiently stabilizing CH−π interactions were additionally found to compensate for the loss of the key cross-ring hydrogen bonds when the D-Pro-L-Pro monomer was utilized (Figure 5). Each of these highlighted examples supports the notion of an interwoven network of energetically balanced non-covalent interactions that can adjust their structure to compensate for mutational weaknesses created in the structure.

Most illustrative of this concept is the X-ray analysis of the catenane $^{F}3s^{Ph}$ ($^{F}3d^{Ph}$ in ref 3) obtained from the combination of the D-Pro-Ac$_4$c and L-^{F}Pro-L-PheGly monomers. A comparative analysis of this structure and the parent $^{H}3a$ indicates that conversion of the naphthyl to a phenyl ring in the CH−π−CH sandwich causes the expected increase[43] in the CH−π distances due to a weakening of the π donor. In response, the network of non-covalent interactions shortens the adjacent inter-ring hydrogen bonds, which concomitantly decreases the width of the aryl barrel and brings the β-1 loops sufficiently close that one of them flips from a type II′ turn to a type I′ to enable a new inter-ring hydrogen bond. These compensating responses from the non-covalent network help to recover some of the loss in stability that the naphthyl-to-phenyl conversion would have otherwise created in the CH−π−CH sandwich.

The cooperativity of the non-covalent network can be further supported by a Gibbs free energy analysis. Since the number of potential n-mers is the same for each DSA reaction, one can semi-quantitatively ascribe an energy of the catenane relative to the ensemble of competing structures by defining an equilibrium between all competitors and the catenanes (eq 1). The resulting equilibrium constant (eq 2) can then be used

$$\text{oligomers(cyclic, acyclic)} \underset{}{\overset{K_{eq}}{\rightleftharpoons}} \text{[2]-catenanes} \qquad (1)$$
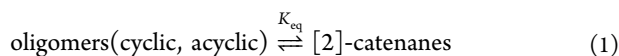
to provide a baseline energy, perturbations from which will generate a $\Delta\Delta G$ that correlates to the perturbation gain or loss in stability.[44] A comparison of these perturbation free energy changes to theoretical measures of the expected gain or loss of stability provides an assessment of the independence of the non-covalent interactions in question. Of particular interest are theoretical measures of CH−π interaction strengths: the computational interaction strength for the CH$_4$---C$_6$H$_6$ complex is 1.45 kcal/mol, while it is 1.92 kcal/mol for the CH$_4$---C$_{10}$H$_8$ complex.[12f]

$$K_{eq} = \frac{[\text{[2]-catenanes}]_{eq}}{[\sum \text{non-catenated oligomers}]_{eq}} \qquad (2)$$

Using the methyl deletion experiments outlined in Scheme 2b and reported in Table 1, one can predict how much of a drop in stability to expect on changing a naphthyl to a phenyl, or removing a CH$_3$−π interaction. Shown in Figure 9 are the relative energies of these same changes determined experimentally using the approach in eqs 1 and 2. In the case of the two D-Ala scenarios, these concentrations were estimates from
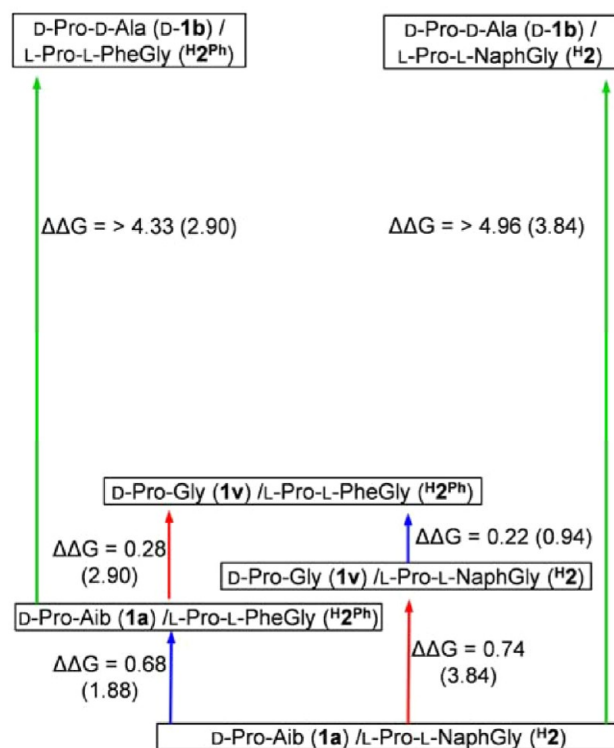
Figure 9. Diagram showing $\Delta\Delta G$ values (kcal/mol) between the six DSAs. The values in parentheses are obtained from the CH−π interaction values reported in the literature.[12f] In each DSA, ΔG was obtained from the concentration ratio of catenanes and non-catenated species, and $\Delta\Delta G$ is the difference of two DSAs' ΔG values. See the details in SI1 Table 4.

the detection limit of the LC-UV traces. Thus, the D-Ala for Aib substitution corresponds to the top right catenane and the lowest energy structure (right green arrow in Figure 9). The measured minimum energy difference is ∼5 kcal/mol. Since each CH$_4$---C$_{10}$H$_8$ interaction is worth 1.92 kcal/mol and two are lost in the D-Ala for Aib mutation, a computational prediction based on CH−π contributions only would predict a $\Delta\Delta G$ of 3.84 kcal/mol. A similar analysis follows for the phenyl series (left green arrow in Figure 9), except that two CH$_4$---C$_6$H$_6$ are worth 2.90 kcal/mol, while the minimum measured energy difference is ∼4.3 kcal/mol. The loss of CH−π interactions is clearly insufficient to describe the phenomenon. As discussed previously, this larger difference is ascribed to the change in intrinsic β-turn preference for these two dipeptides and the absolute requirement for a type II′ turn at the β-2 position. Additionally, contributions no doubt come from the loss of the aforementioned inter-ring hydrogen bonds and the turn-on of the carbonyl−carbonyl repulsions at the i+1 position.

Results at the opposite end of the spectrum were also seen in this data set. For example, when the **Me1** and **Me2** positions (Scheme 2b) are each replaced with hydrogens, one should lose the equivalent of two CH$_4$---C$_{10}$H$_8$ interactions, i.e., 3.84 kcal/mol of stability (recall the β-turn preferences are the same for D-Pro-Aib and D-Pro-Gly). Instead, a $\Delta\Delta G$ of only 0.74 kcal/mol is observed (right red arrow in Figure 9). If one assumes the energy difference reflects a complete loss of the CH−π interactions, this measured difference is most reasonably interpreted as resulting from compensating forces. That is, the network is sufficiently entwined that losses in one part of

the web can be neutralized by gains in another. Phenyl for naphthyl changes were found to behave similarly (blue arrows in Figure 9).

The complexity in the [2]-catenane leads to numerous outcomes on mutagenesis. In some instances, the deletion of seemingly important interactions (naphthyl to phenyl, Figure 9) causes little stability loss, as the network of stabilizing forces can adjust the balance to compensate for the loss. In other circumstances, the excision of a contributing interaction (removal of CH_3 in Scheme 2b) leads to disproportionate losses in stability as larger primary forces are revealed. The question of cooperativity and multi-valency is thus distorted by the situation-dependent strength of the individual non-covalent contributors to stability. That is, in a complex system, the ground-state structure is a compromise (at the individual interaction level) that maximizes the global balance of forces. When the balance is shifted by a mutation, it is possible that a new minimum can be achieved through a rearrangement that best rebalances the energetic contributors.

The parallels to biological scenarios are many. For example, a glutamine (Gln) for glutamic acid (Glu) mutation at the β-6 position of hemoglobin (Hb Machida)[45] causes no conformational or functional changes. The replacement of the same residue by lysine (Lys) or alanine (Ala) also leads to no functional changes (Hb C and Hb G Makassar).[46] By contrast, replacement of this glutamic acid by valine (Hb S) causes a severe structural change and a great diminishment of the deoxygenated Hb S solubility. The subsequent polymerization leads to sickle cell anemia.[47] Thus, several mutations can be accommodated, but one particular mutation leads to a shift toward a different thermodynamic state with negative consequences. The catenane mutagenesis recapitulates this phenomenon.

This [2]-catenane is an uncommon example of complex folding phenomena in a biomimetic system. Rather than isolating a single non-covalent interaction, as is the case in most model systems, these catenanes provide a model for how multivalent systems can compensate for unfavorable mutations and minimize the "toxicity" of many,[48] but not all mutations. This is, of course, a crucial molecular-level characteristic that is required for evolution.

CONCLUSION

Mutation studies of the key non-covalent interactions in the [2]-catenanes described here suggest that the structures are complex enough to mimic numerous key properties of large protein systems. The manipulation of structure and concomitant changes in stability provide insight into the interplay between the components of a network of non-covalent interactions. Even in a self-assembled complex, the complexity of the structure requires that a balance of forces be found, and this balance can fluctuate as structural mutations occur. This network of forces provides the means to compensate for non-optimal changes or even explicitly reject these mutations that create insurmountable energetic costs. These fundamental properties are therefore not exclusively found in large biological systems.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Synthesis of monomers and their physical data, HPLC traces including oligomer speciation, LC-MS, LC-TOF analyses of various DSAs, yields of [2]-catenanes in the additional DSA

solutions, and calculations of the $\Delta\Delta G$ values between the six DSAs using the oligomer speciation. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

mgagne@unc.edu; mlwaters@unc.edu

Notes

The authors declare no competing financial interest.

REFERENCES

(1) (a) Hunter, C. A.; Anderson, H. L. *Angew. Chem., Int. Ed.* **2009**, *48*, 7488−7499. (b) Whitty, A. *Nat. Chem. Biol.* **2008**, *4*, 435−439. (c) Badjic, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Acc. Chem. Res.* **2005**, *38*, 723−732. (d) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 6596−6616. (e) Shinikai, S.; Ikeda, M.; Sugasaki, A.; Takeuchi, M. *Acc. Chem. Res.* **2001**, *34*, 494−503. (f) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754−2794.

(2) Ercolani, G. *J. Am. Chem. Soc.* **2003**, *125*, 16097−16103.

(3) Chung, M.-K.; White, P. S.; Lee, S. J.; Waters, M. L.; Gagne, M. R. *J. Am. Chem. Soc.* **2012**, DOI: 10.1021/ja302345n.

(4) (a) Nishio, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13873−13900 and references therein. (b) Takahashi, O.; Kohno, Y.; Nishio, M. *Chem. Rev.* **2010**, *110*, 6049−6076 and references therein.

(5) (a) Carrillo, R.; López-Rodríguez, M.; Martín, V. S.; Martín, T. *Angew. Chem., Int. Ed.* **2009**, *48*, 7803−7809. (b) He, Q.; Rohani, S.; Zhu, J.; Gomaa, H. *Chirality* **2012**, *24*, 119−128. (c) Saigo, K.; Kobayashi, Y. *Chem. Rec.* **2007**, *7*, 47−56 and references therein. (d) Suezawa, H.; Ishihara, S.; Umezawa, Y.; Tsuboyama, S.; Nishio, M. *Eur. J. Org. Chem.* **2004**, *23*, 4816−4822. (e) Kobayashi, Y.; Kurasawa, T.; Kinbara, K.; Saigo, K. *J. Org. Chem.* **2004**, *69*, 7436−7441.

(6) (a) Nishio, M.; Umezawa, Y.; Honda, K.; Tsuboyama, S.; Suezawa, H. *CrystEngComm* **2009**, *11*, 1757−1788 and references therein. (b) Schneider, H.-J. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924−3977 and references therein.

(7) (a) Scarso, A.; Trembleau, L.; Rebek, J., Jr. *J. Am. Chem. Soc.* **2004**, *126*, 13512−13518. (b) Bhayana, B.; Wilcox, C. S. *Angew. Chem., Int. Ed.* **2007**, *46*, 6833−6836. (c) For recognition involving carbohydrates, see ref 4a and references therein.

(8) Nishio, M. *Tetrahedron* **2005**, *61*, 6923−6950 and references therein.

(9) Nishio, M. *CrystEngComm* **2004**, *6*, 130−158 and references therein.

(10) (a) Ganguly, H. K.; Majumder, B.; Chattopadhyay, S.; Chakrabarti, P.; Basu, G. *J. Am. Chem. Soc.* **2012**, *134*, 4661−4669. (b) Wilger, D. J.; Park, J. H.; Hughes, R. M.; Cuellar, M. E.; Waters, M. L. *Angew. Chem., Int. Ed.* **2011**, *50*, 12201−12204.

(11) Chung, M.-K.; White, P. S.; Lee, S. J.; Gagné, M. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 8683−8686.

(12) (a) Thomas, K. M.; Naduthambi, D.; Zondlo, N. J. *J. Am. Chem. Soc.* **2006**, *128*, 2216−2217. (b) Carroll, W. R.; Zhao, C.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D. *Org. Lett.* **2011**, *13*, 4320−4323. (c) Kim, E.-I.; Paliwal, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 11192−11193. (d) Fujii, A.; Hayashi, H.; Park, J. W.; Kazama, T.; Mikami, N.; Tsuzuki, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14131−14141. (e) Nilsson Lill, S. O. *J. Mol. Graphics Modell.* **2010**, *29*, 178−187. (f) Tsuzuki, S.; Fujii, A. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2584−2594. (g) Gil, A.; Branchadell, V.; Bertran, J.; Oliva, A. *J. Phys. Chem. B* **2007**, *111*, 9372−9379.

(13) (a) Das, C.; Naganagowda, G. A.; Karle, I. L.; Balaram, P. *Biopolymers* **2001**, *58*, 335−346. (b) Raghothama, S. R.; Awasthi, S. K.;

Balaram., P. *J. Chem. Soc., Perkin Trans. 2* 1998, 137−143. (c) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* 1997, 119, 2303−2304. (d) Kantharaju; Raghothama, S.; Aravinda, S.; Shamala, N.; Balaram, P. *Biopolymers (Pept. Sci.)* 2010, 94, 360−370. (e) Aravinda, S.; Harini, V. V.; Shamala, N.; Das, C.; Balaram, P. *Biochemistry* 2004, 43, 1832−1846.

(14) (a) Lai, J. R.; Huck, B. R.; Weisblum, B.; Gellman, S. H. *Biochemistry* 2002, 41, 12835−12842. (b) Paul, P. K. C.; Sukumar, M.; Bardi, R.; Piazzesi, A. M.; Valle, G.; Toniolo, C.; Balaram, P. *J. Am. Chem. Soc.* 1986, 108, 6363−6370. (c) Harini, V. V.; Aravinda, S.; Rai, R.; Shamala, N.; Balaram, P. *Chem.—Eur. J.* 2005, 11, 3609−3620. (d) Benedetti, E. *Biopolymers (Pept. Sci.)* 1996, 40, 3−44. (e) Toniolo, C.; Crisma, M.; Formaggio, F.; Peggion, C. *Biopolymers (Pept. Sci.)* 2001, 60, 396−419.

(15) 7−10 days for DSAs containing $^{H}2$ or $^{F}2$, 2−3 days for DSAs using $^{HO}2$ or $^{HO}2^{Ph}$, 3−4 days for DSAs having $^{H}2^{Ph}$ or $^{F}2^{Ph}$, and 3−5 days for DSAs having $^{H}2^{Ph}_{Aze}$ or $^{F}2^{Ph}_{Aze}$.

(16) The standard dihedral angle comparisons of the type II′ and I′ turns show almost 180° differences in Ψ_{i+1} and Φ_{i+2} values, which correspond to the turn backbone carbonyl (i+1) and amide (i+2) groups: type II′, Φ_{i+1}, 60°, Ψ_{i+1}, −126°, Φ_{i+2}, −91°, and Ψ_{i+2}, 1°; type I′, Φ_{i+1}, 55°, Ψ_{i+1}, 38°, Φ_{i+2}, 78°, and Ψ_{i+2}, 6°.

(17) Only a trace amount of [2]-catenanes ($\ll$1%) is generated. The loss was calculated from the HPLC-UV trace total area (289 nm) comparison of [2]-catenanes generated from a DSA reaction using a D-Pro-Aib (D-**1a**) and a L-Pro-L-NaphGly (L,L-^{H}**2**).

(18) (a) McComas, C. C.; Crowley, B. M.; Boger, D. L. *J. Am. Chem. Soc.* 2003, 125, 9314−9315. (b) Walsh, C. T.; Fisher, S. L.; Park, I.-S.; Prahalad, M.; Wu, Z. *Chem. Biol.* 1996, 3, 21−28. (c) Walsh, C. T. *Science* 1993, 261, 308−309. (d) Bugg, T. D. H.; Wright, G. D.; Dutka-Malen, S.; Arthur, M.; Courvalin, P.; Walsh, C. T. *Biochemistry* 1991, 30, 10408−10415.

(19) Since the [2]-catenanes are D-component-rich, the residual oligomers populate the more L-component-rich species with D-**1a**. Despite this bias, the *n*-mer ratios are largely unaffected.

(20) An equivalent logical analysis can be constructed for the **4** catenanes.

(21) Since [2]-catenane **4** contains three units of L,L-^{H}**2**, a separate but logically equivalent set of isomers should result.

(22) Similarly consistent with case 2 was the detection (by LC-TOF) of only two isomers of **3** with 1 equiv of D-Pro-D-Ala (Figure 3a; see additional details in SI1 Figure 7).

(23) Gibbs, A. C.; Bjorndahl, T. C.; Hodges, R. S.; Wishart, D. S. *J. Am. Chem. Soc.* 2002, 124, 1203−1213.

(24) The introduction of a CF$_3$ at Me **1** (Scheme 2b) (D-Pro-3,3,3-trifluoro-L-Ala) completely destabilized the catenanes, and no catenane formation was observed.

(25) The Hunter−Sanders model predicts an increased π/π repulsion with electron-donating substituents, indicating that electrostatic effects alone do not explain the observed stability trend. X-ray analysis of ^{HO}**3n** (^{HO}**3c** in ref 3) did not reveal any close contacts to the *para*-substituent. Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* 1990, 112, 5525−5534.

(26) X-ray structure analysis of the catenane ^{F}**3s**^{Ph} (^{F}**3d**^{Ph} in ref 3) containing the β-turns of Ac$_4$c residues also shows that these β-turns are quite similar to those of Aib. Previous studies have shown Ac$_3$c residues to form the similar (but more rigid) β-turns: (a) Alemán, C. *J. Phys. Chem. B* 1997, 101, 5046−5050. (b) Gatos, M.; Formaggio, F.; Crisma, M.; Toniolo, C.; Bonora, G. M.; Benedetti, Z.; Di Blasio, B.; Iacovino, R.; Santini, A.; Saviano, M.; Kamphuis, J. *J. Pept. Sci.* 1997, 3, 110−122. (c) Crisma, M.; Bonora, G. M.; Toniolo, C.; Barone, V.; Benedetti, Z.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A.; Fraternali, F.; Bavoso, A.; Lelj, F. *Int. J. Biol. Macromol.* 1989, 11, 345−352. (d) Toniolo, C. *Biopolymers* 1989, 28, 247−257.

(27) (a) Ballano, G.; Zanuy, D.; Jiménez, A. I.; Cativiela, C.; Nussinov, R.; Alemán, C. *J. Phys. Chem. B* 2008, 112, 13101−13115. (b) Zanuy, D.; Rodrígues-Ropero, F.; Haspel, N.; Zheng, J.; Nussinov, R.; Alemán, C. *Biomacromolecules* 2007, 8, 335−3146. (c) Zheng, J.;

Zanuy, D.; Haspel, N.; Tsai, C.-J.; Alemán, C.; Nussinov, R. *Biochemistry* 2007, 46, 1205−1218.

(28) Even when a weaker π-donor, L-Pro-L-PheGly, was employed, significant amounts of [2]-catenanes were generated (SI1 Table 2).

(29) In a three-component DSA using D-Pro-D-Ala/D-Pro-Gly (1:1) as the D-monomers with L,L-^{H}**2**, the catenane compositions indicate the complete interchangeability between D-Pro-Aib and D-Pro-Gly (see SI1 Figures 11 and 12).

(30) (a) Chatterjee, B.; Saha, I.; Raghothama, S.; Aravinda, S.; Rai, R.; Shamala, N.; Balaram, P. *Chem.—Eur. J.* 2008, 14, 6192−6204. (b) Rai, R.; Raghothama, S.; Balaram, P. *J. Am. Chem. Soc.* 2006, 128, 2675−2681. (c) Nair, M.; Vijayan, M.; Venkatachalapathi, Y. V.; Balaram, P. *J. Chem. Soc., Chem. Commun.* 1979, 1183−1184.

(31) (a) Park, H. S.; Kim, C.; Kang, Y. K. *Biophys. Chem.* 2003, 105, 89−104. (b) Antohi, O.; Sapse, A.-M. *J. Mol. Struct. (Theochem)* 1998, 430, 247−258. (c) Aubry, A.; Cung, M. T.; Marraud, M. *J. Am. Chem. Soc.* 1985, 107, 7640−7647.

(32) The standard dihedral angle comparison of these types and type VIII shows more than 110° differences in Φ_{i+1}, Ψ_{i+1}, and Ψ_{i+2} values, which relate to the turn backbone amide and carbonyl (i+1) and carbonyl (i+2) groups: type II, Φ_{i+1}, −60°, Ψ_{i+1}, 131°, Φ_{i+2}, 84°, and Ψ_{i+2}, 1°; type I′, Φ_{i+1}, 55°, Ψ_{i+1}, 38°, Φ_{i+2}, 78°, and Ψ_{i+2}, 6°; type II′, Φ_{i+1}, 60°, Ψ_{i+1}, −126°, Φ_{i+2}, −91°, and Ψ_{i+2}, 1°.

(33) (a) Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/π Interaction: Evidence, Nature, and Consequences*; Wiley-VCH; New York, 1998. (b) Fujii, A.; Shibasaki, K.; Kazama, T.; Itaya, R.; Mikami, N.; Tsuzuki, S. *Phys. Chem. Chem. Phys.* 2008, 10, 2836−2843. (c) Gotch, A. J.; Pribble, R. N.; Ensminger, F. A.; Zwier, T. S. *Laser Chem.* 1994, 13, 187−205. (d) Gord, J. R.; Garrett, A. W.; Bandy, R. E.; Zwier, T. S. *Chem. Phys. Lett.* 1990, 171, 443−450.

(34) L,L-$^{cis-F}$**2** and L,L-$^{cis-HO}$**2**^{Ph} monomers could not be separated from their epimers. Therefore, the experiments using these monomers were omitted.

(35) See SI2 for HPLC traces of DSA solutions and LC-TOF analyses of the generated catenanes.

(36) (a) Dey, R. C.; Seal, P.; Chakrabarti, S. *J. Phys. Chem. A* 2009, 113, 10113−10118. (b) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* 2000, 122, 3746−3753.

(37) (a) Kim, W.; Hardcastle, K. I.; Conticello, V. P. *Angew. Chem., Int. Ed.* 2006, 45, 8141−8145. (b) Lesarri, A.; Cocinero, E. J.; López, J. C.; Alonso, J. L. *J. Am. Chem. Soc.* 2005, 127, 2572−2579. (c) Thomas, K. M.; Naduthambi, D.; Tririya, G.; Zondlo, N. J. *Org. Lett.* 2005, 7, 2397−2400. (d) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. K. *J. Am. Chem. Soc.* 2002, 124, 2497−2505. (e) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem., Int. Ed.* 2001, 40, 923−925.

(38) (a) Okuyama, K.; Miyama, K.; Morimoto, T.; Masakiyo, K.; Mizuno, K.; Bächinger, H. P. *Biopolymers* 2011, 95, 628−640. (b) Shoulders, M. D.; Satyshur, K. A.; Forest, K. T.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* 2010, 107, 559−564. (c) Jiravanichanun, N.; Nishino, N.; Okuyama, K. *Biopolymers* 2006, 81, 225−233.

(39) The fact that the *cis*-4-F-L-Pro-Gly segment has been reported to more strongly favor the type I β-turn than *trans*-4-F-L-Pro-Gly may attenuate this possibility: Kim, W.; McMillan, A.; Snyder, J. P.; Conticello, V. P. *J. Am. Chem. Soc.* 2005, 127, 18121−18132.

(40) (a) Tsuzuki, S.; Honda, K.; Fujii, A.; Uchimaru, T.; Mikami, M. *Phys. Chem. Chem. Phys.* 2008, 10, 2860−2865. (b) Morita, S.; Fujii, A.; Mikami, N.; Tsuzuki, S. *J. Phys. Chem. A* 2006, 110, 10583−10590. (c) Ringer, A. L.; Figgs, M. S.; Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* 2006, 110, 10822−10828.

(41) Ran, J.; Wong, M. W. *J. Phys. Chem. A* 2006, 110, 9702−9709.

(42) (a) Blessing, R. H.; Smith, D. *Acta Crystallogr.* 1982, B38, 1203−1207. (b) Berman, H. M.; McGandy, E. L.; Burgner, J. W., II; VanEtten, R. L. *J. Am. Chem. Soc.* 1969, 91, 6177−6182. (c) Jhon, J. S.; Kang, Y. K. *J. Phys. Chem. B* 2007, 111, 3496−3507.

(43) Computational studies show that the optimum aromatic−CH intermolecular separation (R) increase from the pyrene−methane ($R = 3.4$ Å) to the naphthalene−methane ($R = 3.6$ Å) to the benzene−

methane (R = 3.8 Å) clusters, respectively. R corresponds to the distance between the aromatic plane and the methane carbon atom. See refs 40a and 12f.

(44) In practice this was accomplished by carrying out a full speciation analysis and expressing the catenane concentration over the sum of the other oligomer concentrations (see SI1 Table 4).

(45) Harano, T.; Harano, K.; Ueda, S.; Shibata, S.; Imai, K.; Seki, M. *Hemoglobin* **1982**, *6*, 531−535.

(46) (a) Hunt, J. A.; Ingram, V. M. *Biochim. Biophys. Acta* **1960**, *12*, 409−421. (b) Blackwell, R. Q.; Oemijati, S.; Pribadi, W.; Weng, M.-I.; Liu, C.-S. *Biochim. Biophys. Acta* **1970**, *214*, 396−401.

(47) (a) Ingram, V. M. *Nature* **1957**, *180*, 326−328. (b) Ingram, V. M. *Biochim. Biophys. Acta* **1959**, *36*, 402−411. (c) Bihoreau, M. T.; Baudin, V.; Marden, M.; Lacaze, N.; Bohn, B.; Kister, J.; Schaad, O.; Dumoulin, A.; Edelstein, S. J.; Poyart, C.; Pagnier, J. *Protein Sci.* **1992**, *1*, 145−150. (d) Baudin-Chich, V.; Pagnier, J.; Marden, M.; Bohn, B.; Lacaze, N.; Kister, J.; Schaad, O.; Edelstein, S. J.; Poyart, C. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1845−1849. (e) Adachi, K.; Kim, J.; Travitz, R.; Harano, T.; Asakura, T. *J. Biol. Chem.* **1987**, *262*, 12920−12925. (f) Padlan, E. A.; Love, W. E. *J. Biol. Chem.* **1985**, *260*, 8272−8279. (g) Magdoff-Fairchild, B.; Poillon, W. N.; Li, T.-I.; Bertles, J. F. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 990−994. (h) Harrington, D. J.; Adachi, K.; Royer, W. E., Jr. *J. Mol. Biol.* **1997**, *272*, 398−407. (i) Nagel, R. L.; Johnson, J.; Bookchin, R. M.; Garel, M. C.; Rosa, J.; Schiliro, G.; Wajcman, H.; Labie, D.; Moo-Penn, W.; Castro, O. *Nature* **1980**, *283*, 832−834.

(48) The only two fully "toxic" mutations are the change in the β-turn conformation at β-2 and β-4 and the incorporation of a polar group in a position that will interact with the π-system. Other typical types of mutations, such as changes in steric and electronic interactions at the Aib position, "knock down" the formation of the catenane, but do not completely knock it out.